

WHAT IS CLAIMED IS:

1. A library of dual-domain nucleic acid molecules each of which has
 - (a) a first and a second domain;
 - (b) separating and linking said domains, a linker which is a member of a randomized library of linkers that
 - (i) vary in size and nucleotide sequence,
 - (ii) consist of a repeated pattern of degenerate repeated triplet nucleotides.
2. The library of molecules of claim 1, wherein said repeated pattern of degenerate repeated triplet nucleotides of said linkers having the following properties:
 - (i) position 1 of each repeated triplet cannot be the same nucleotide as position 2 of the repeated triplet; or
 - (ii) position 2 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; or
 - (iii) position 1 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet.
3. The library of molecules of claim 2 wherein the nucleotide in the first and second positions of each repeated triplet is selected from any two of deoxyadenosine, deoxyguanosine, deoxycytidine or deoxythymidine.
4. The library of molecules of claim 3, wherein
 - (i) position 1 of each repeated triplet is deoxyadenosine or deoxyguanosine;
 - (ii) position 2 of each repeated triplet is deoxycytidine or deoxyguanosine; and
 - (iii) position 3 of each repeated triplet is deoxythymidine.
5. The library of molecules of claim 1 wherein at least one of said domains binds to a protein.
6. The library of molecules of claim 5 wherein both of said domains bind to a protein.

7. The library of molecules of claim 1 wherein at least one of said domains binds to a nucleic acid that is not a member of said library.

5 8. The library of molecules of claim 7 wherein both of said domains bind to a nucleic acid that is not a member of said library.

9. The library of molecules of any of claims 1-4 wherein said first and said second domains are coding sequences.

10 10. The library of molecules of any of claims 1-8 produced in plant cells.

11. The library of molecules of claim 9 produced in plant cells.

15 12. A dual-domain nucleic acid molecule selected from the library of any of claims 1-8.

13. A dual-domain nucleic acid molecule selected from the library of claim 9.

20 14. A dual-domain nucleic acid molecule selected from the library of claim 10.

15. A dual-domain nucleic acid molecule selected from the library of claim 11.

16. A library of dual-domain polypeptide molecules each of which is described by the formula D₁-L -D₂ wherein

25 (a) D₁ and D₂ are polypeptide domains and

(b) L is a peptide or polypeptide linker which is a member of a randomized library of linkers that vary in size and sequence, which library is encoded by nucleic acid sequences consisting of a repeated pattern of degenerate repeated triplet nucleotides.

17. A library of multi-domain polypeptide molecules each of which comprises polypeptide domains **D** each pair of which is linked by a peptide or polypeptide linker **L**, each molecule being described by the formula **D_xL_y** wherein

5 x is an integer between 2 and 20,

y is an integer between 1 and 19, with the proviso that for any value of x, y=x-1;

10 **D₁** is bonded to a single C-terminal linker;

the C-terminal-most **D** is bonded to a single N-terminal linker;

each of **D₂** to **D₁₉** are bonded to a N-terminal and a C-terminal linker;

each **L** is a member of a randomized library of linkers that vary in size and sequence,

10 said linker library being encoded by nucleic acid sequences consisting of a repeated pattern of degenerate repeated triplet nucleotides.

15 18. The library of dual domain polypeptide molecules of claim 16, or multi-domain polypeptide molecules of claim 17, wherein each linker in said library

- (i) has a length of between about one and 50 amino acid residues
- (ii) between 1 and about 20 different amino acids wherein each repeated pattern of degenerate triplet bases encodes between 1 and about 12 different amino acids.

20 19. The library of polypeptide molecules of claim 18, wherein said repeated pattern of degenerate repeated triplet nucleotides encoding said linkers having the following properties:

- (i) position 1 of each repeated triplet cannot be the same nucleotide as position 2 of the repeated triplet; or
- (ii) position 2 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; or
- (iii) position 1 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet.

30

002260 002261 002262 002263 002264 002265 002266 002267 002268 002269 002270

20. The library of polypeptide molecules of claim 19 wherein the nucleotide in the first and second positions of each repeated triplet is selected from any two of deoxyadenosine, deoxyguanosine, deoxycytidine or deoxythymidine.

5 21. The library of polypeptide molecules of claim 20, wherein
(i) position 1 of each repeated triplet is deoxyadenosine or deoxyguanosine;
(ii) position 2 of each repeated triplet is deoxycytidine or deoxyguanosine; and
(iii) position 3 of each repeated triplet is deoxythymidine.

10 22. The library of dual-domain polypeptide molecules of claim 16 or multi-domain polypeptide molecules of claim 17 produced in plant cells.

23. The library of polypeptide molecules of claim 18 produced in plant cells.

24. The library of polypeptide molecules of claim 19 produced in plant cells.

15 25. The library of polypeptide molecules of claim 20 produced in plant cells.

26. The library of polypeptide molecules of claim 21 produced in plant cells.

27. A dual-domain polypeptide molecule selected from the library of claim 16.

16 28. A multi-domain polypeptide molecule selected from the library of claim 17.

29. A dual domain polypeptide molecule or multi-domain polypeptide molecule selected from the library of claim 18.

20 30. A dual domain polypeptide molecule or multi-domain polypeptide molecule selected from the library of claim 19.

31. A dual domain polypeptide molecule or multi-domain polypeptide molecule selected from the library of claim 20.

32. A dual domain polypeptide molecule or multi-domain polypeptide molecule selected from the library of claim 21.

33. A three domain peptide selected from the library of claim 17

which is a dual domain scFv polypeptide linked to a third polypeptide domain.

34. The three domain polypeptide of claim 33 wherein the third domain is a toxin polypeptide or an enzyme.

5 35. A method of generating the library of dual-domain nucleic acids of
claim 1, comprising:

- a. obtaining two template DNA sequences that comprises the first and the second domains;
- b. preparing amplification primer pairs which amplify the first and second domains where each primer pair comprises an upstream primer and a downstream primer, each primer having a 5' end and a 3' end, wherein the downstream primer for the first domain or the upstream primer for the second domain comprises a nontemplated sequence,
 - said nontemplated sequence comprising a repeated pattern of degenerate repeated triplet nucleotides,
 - wherein at least two of the 5' terminal triplets of said repeated pattern of degenerate repeated triplet nucleotides have the same degenerate sequence;
- c. amplifying the domains with the amplification primers to generate at least one population of nucleic acid domains having different lengths and sequences in the non-templated sequence; and
- d. ligating the nucleic acid domains generated in step (c) to generate said a population of dual-domain molecules.

36. The method of claim 35, wherein said repeated pattern of degenerate repeated triplet nucleotides in at least one of said primers has the following properties:

- (i) position 1 of each repeated triplet cannot be the same nucleotide as position 2 of the repeated triplet; or
- 5 (ii) position 2 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; or
- (iii) position 1 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet.

10 37. The method of claim 35 wherein at least one of the primers contains a non-templated endonuclease recognition site.

38. The method of claim 35 wherein said template DNA sequences are made by reverse transcription of mRNA.

15 39. The method of claim 35 further comprising the step of ligating the population of dual-domain nucleic acids to vectors.

40. The method of claim 39, further comprising the step of introducing said vector into a host.

20 41. The method of claim 40 wherein said nucleic acid domains encode polypeptide domains, and which method further comprises the step of expressing dual-domain polypeptides encoded by said dual-domain nucleic acids.

42. The method of claim 39 wherein further comprising the step of transcribing RNA from said vectors

25 43. The method of claim 42 wherein said vectors are compatible with replication and/or expression of said nucleic acids in plant cells, said method further comprising the steps of introducing the transcribed said RNA into a plant cell and expressing the dual-domain polypeptide.

44. A population of dual-domain polypeptides or a dual-domain polypeptide selected therefrom, produced by the method of claim 41.

45. A population of dual-domain polypeptides or a dual-domain polypeptide selected therefrom, produced in plant cells by the method of claim 43.

46. A method of producing the polypeptide of claim 27 comprising the steps of:

- 5 (a) joining a nucleic acid encoding the first domain of the polypeptide to a nucleic acid encoding a first part of a linker to produce a first nucleic acid construct;
- (b) joining the nucleic acid encoding a second part of the linker to a nucleic acid encoding the second domain of the polypeptide to produce a second nucleic acid construct;
- 10 (c) incorporated said first and said second constructs into a transient plant expression vector in frame so that, when expressed, the polypeptide bears the first and second domain separated by the linker as described by the formula D₁-L -D₂.
- 15 (d) transfecting a plant with the vector so that the plant transiently produces the polypeptide; and
- (e) recovering the polypeptide as a soluble, functionally-folded protein.

47. The method of claim 46 wherein the plant is a plant cell.

48. A linker nucleic acid molecule or sequence that joins two nucleic acid domains or two nucleic acid sequences encoding two polypeptide domains, which has a pattern of degenerate repeated triplet nucleotides with the following properties:

- 20 (i) position 1 of each repeated triplet cannot be the same nucleotide as position 2 of the repeated triplet; or
- (ii) position 2 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; or
- 25 (iii) position 1 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; and
- (iv) wherein said molecule or sequence that joins said domains does not encode Gly₄Ser or a repeat thereof.

49. A library of linker nucleic acid molecules or sequences each of which joins two nucleic acid domains or two nucleic acid sequences encoding two polypeptide domains, each of which has a pattern of degenerate repeated triplet nucleotides with the following properties:

- 5 (i) position 1 of each repeated triplet cannot be the same nucleotide as position 2 of the repeated triplet; or
- (ii) position 2 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; or
- (iii) position 1 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; and
- 10 (iv) wherein each of said molecules or sequences that joins said domains does not encode Gly₄Ser or a repeat thereof.

00020600-2020-00000

50. A method for making the library of linker nucleic acid molecules or sequences of claim 49, comprising:

- (a) obtaining two template DNA sequences that comprise the first and the second domains;
- 5 (b) preparing amplification primer pairs which amplify the first and second domains where each primer pair comprises an upstream primer and a downstream primer, each primer having a 5' end and a 3' end, wherein the downstream primer for the first domain or the upstream primer for the second domain comprises a nontemplated sequence,
 - 10 said nontemplated sequence comprising said repeated pattern of degenerate repeated triplet nucleotides, wherein at least two of the 5' terminal triplets of said repeated pattern of degenerate repeated triplet nucleotides have the same degenerate sequence;
- (c) amplifying the domains with the amplification primers to generate at least one population of nucleic acid domains having different lengths and sequences in the non-templated sequence; and
- 15 (d) ligating the nucleic acid domains generated in step (c) to generate said population of dual-domain molecules.
- (e) excising or amplifying said linker nucleic acid molecules or sequences 20 from said population of dual domain molecules.

00000000000000000000000000000000

51. A method for making a linker nucleic acid molecule or sequence that joins two nucleic acid domains or two nucleic acid sequences encoding two polypeptide domains, which has a pattern of degenerate repeated triplet nucleotides with the following properties:

- (i) position 1 of each repeated triplet cannot be the same nucleotide as position 2 of the repeated triplet; or
- (ii) position 2 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; or
- (iii) position 1 of each repeated triplet cannot be the same nucleotide as position 3 of the repeated triplet; and
- (iv) wherein said molecule or sequence that joins said domains does not encode Gly₄Ser or a repeat thereof.

said method comprising the steps of:

- (a) making the library of linker nucleic acid molecules or sequences in accordance with the method of claim 49
- (b) selecting and isolating said linker molecule or sequence from said library.